REMARKS

As noted above, the Applicant appreciates the Examiner's thorough examination of the subject application.

Claims 1-5, 7-28 and 30-38 remain in the application. Claims 1-5, 7-28 and 30-38 were rejected in the Office Action mailed September 14, 2005. Claims 1, 14, and 24 have been amended in order to more clearly define Applicant's invention. No new matter has been added.

Applicant respectfully traverses the rejection of the claims based on the foregoing amendments and the following remarks.

Examiner's Response to Amendment

Concerning item 1 of the Office Action, the Applicant thanks the Examiner for entering the previous amendment to the claims (i.e., Amendment "B").

Examiner's Response to Arguments

Concerning item 2 of the Office Action, the Examiner, after correctly admitting that U.S. Patent No. 5,615,296 to Stanford at al. ("Stanford") does not disclose "semantic processing," stated that in was unclear to the Examiner what the Applicant's use of the term "semantic processing" included. The Examiner subsequently offered an interpretation of the term "semantic processing."

The Applicant respectfully takes issue with and traverses. the Examiner's general characterization of the claimed term "semantic representation" including the specific suggestion that the semantic representation "is nothing more than the command representation used by the computer (e.g. a call to a computer software function in computer code) that corresponds with the user's verbal input." In addition, the Applicant respectfully takes issue with and traverses the Examiner's interpreted definition that the term simply means "a representation of the meaning of the user's verbal input."

As one skilled in the art would understand from the present application and related application serial no. 09/815,769 (now U.S. Patent No. 6,895,377 to Kroeker et al. ("Kroeker"), which is incorporated into the present application by reference, a "semantic representation" includes "semantic information (e.g., a semantic description)". *See* Kroeker, col. 4, line 4. The present application clearly equates such semantic data with the term "semantic representation". *See*, specification as filed, page 14, line 28 – page 15, line 2. The Kroeker patent teaches that a semantic description:

> includes a list of semantic attributes. Each semantic attribute may be a value, a category, an operator, or a tree of such things. Attribute values are specific items, such as the number 3, that have meaning when interpreted at run time. Categories are symbols, possibly with values, that mark the path for future semantic interpretation. Operators control the combination of class instances and provide powerful, extensible, and general techniques for semantic evaluation. Note that any given class may have, and may be interpreted in accordance with, multiple categories. These categories control different semantic interpretations of the same class instance. Collectively, the categories describe all possible valid interpretations of the class. Because all classes are context free, they may be re-used and reinterpreted in accordance with different contexts. For example, a class representing the numbers from 20-99 may be reused in several instances where there is a phonetic input corresponding to a number.
>
> [Kroeker, col. 4, lines 15-33]

The Kroeker patent further teaches a phonetic data processing system and method that, among other things, function to "produce a set of semantic data representing a plurality of valid responses". *See* Kroeker, col. 2, lines 40-41, and col. 6, lines 64-65.

For clarification in light of the Examiner's remarks, independent claims 1, 14, and 28 have been amended to include that the semantic representation, or semantic data message, includes "a plurality of possible valid interpretations of said voice data message".

The Examiner's characterizations of U.S. Patent No. 5,615,296 to Stanford at al. are traversed, as discussed below.

Claim Rejections – 35 U.S.C. § 103

Claims 14-18, 22, and 23

Concerning items 4-5, of the Office Action, claims 14-18, 22, and 23 were rejected under 35 U.S.C. § 103(a) as being unpatentable over U.S. Patent No. 5,615,296 to Stanford et al. ("Stanford"), taken in view of U.S. Patent No. 5,881,230 to Christensen et al. ("Christensen"). The Applicant respectfully traverses this rejection and asks for reconsideration for the following reasons.

One requirement for a rejection under 35 U.S.C. § 103 is that the cited reference(s) teach or suggest all of the limitations of the claims at issue. In this situation, the combination of Stanford and Christensen fails to teach or suggest all of the limitations of claim 14, which is the base claim for claims 15-18, 22, and 23. The claims of the subject application are directed to a remote server object architecture for implementing computing-intensive speech recognition systems that include simultaneous operations of multiple functional units, e.g., a number of discrete servers. The servers can be located in separate computer systems so that simultaneous operations can occur. Four types of functions related to speech recognition have been identified and segregated in different operating units of the claimed architectures and methods.

Claim 14 of the present application (from which claims 15-18, 22, and 23 depend) recites the following, including the four functions and operating units used for the present invention for speech recognition:

> A method of processing speech comprising: (A) receiving, at a first server object, a voice data message from a telephone network; (B) transmitting said voice data message over a first computer network to a second server object; (C) converting said voice data message to a phonetic data message in said second server object; (D) transmitting said phonetic data message from said second server object to a third server object over said first computer network; (E) converting said phonetic data message to a syntactic data message in said third server object; (F) transmitting said syntactic data message from said third server object to a fourth server object over said first computer network; and (G) converting, in said fourth server object, said syntactic data message

> to a semantic data message that includes a plurality of possible valid interpretations of said voice data message.
>
> [Emphasis added]

In contrast, Stanford teaches a system for a continuous speech speaker independent speech recognizer and voice response system for a plurality of user. The system employs high-speed context switching to modify or activate Audio WAV voice response files. The system uses dialogue history to activate selected context. Baukus-Naur Form (BNF) grammars and WAV files provide phrase or sentence long dialogue prompts and audio prompts. *See* Stanford, Abstract.

For the rejection, the Examiner stated that Stanford discloses a method of processing speech including:

> converting said voice data message to a phonetic data message in said second server object (vector quantization block 104 uses Cepstral coefficients converted from the input speech, column 8 line 56 to column 9 line 4; to select the closest codebook values, each codebook representing phonetic data, column 9, lines 38-48).

Applicant respectfully traverses and takes issue with this characterization of Stanford. Applicant notes that the quoted language enclosed by parentheses was supplied by the Examiner and is not actually found in Stanford (while the language preceding the parentheses is taken directly from claim 14). The cited portion of Stanford states in relevant part "For the final step of vector quantization, block 104 refers to a code book derived from the training procedure, just described, to determine which cluster center is closest to the frame Cepstral coefficients." Stanford describes the referenced code book as follows:

> In the training process, a number of sentences are taken, currently between ten to fifteen thousand, and segmented into frames, from which auto-correlations and Cepstrum coefficients are calculated. A clustering procedure is applied to segregate the Cepstrum frame features into two hundred and fifty six classes using a k-means type clustering procedure described in [5]. The centers of these Cepstrum clusters, and their class labels, taken together, are

> hereafter referred to as "code books." [Emphasis added]

In addition the underlined text above, Stanford also teaches that "[t]he adjustment of the training data described . . . are crucial in enabling the base Sphinx recognition system to operate over telephony equipment, and hence to the invention described herein." *See* Stanford, col. 9, lines 26-29. Thus, the code books of Stanford are not phonetic data as claimed by the Applicant. Accordingly, Stanford does not teach or suggest all of the limitations as arranged in claims 14, including "(C) converting said voice data message to a phonetic data message in said second server object," and actually teaches away from the claimed invention.

For this same rejection, the Examiner stated further that Stanford teaches "converting said syntactic data message to a semantic data message, representative of said voice data message in said fourth server object (recognition server converts word sequences to communicate with user applications, column 9, lines 51-53)". Applicant respectfully traverses and takes issue with this characterization of Stanford. Applicant notes that the quoted language enclosed by parentheses was supplied by the Examiner and is not actually found in Stanford (while the language preceding the parentheses is taken directly from claim 14).

The cited portion of Stanford states "The Recognition Server communicates with user applications or Recognition Clients (block 110)." Stanford goes on to explain the operation of its Recognition Server:

> The speech recognition server acquires speech data until a (user adjustable, but most commonly 0.6 seconds) period of silence. Recognition is terminated when this period is observed, and it is assumed that the person is done speaking.
>
> The context voice response files 400 which include a plurality of files 404, are connected to the recognition server 108. When the recognition server designates a new context over the context control line, a new context voice response file 404 is accessed and provide to the voice response output 402. This is carried out in parallel with the invocation of a new recognition context

> in the beam search module 106. As the recognition server 108 receives corresponding text string messages from the beam search module 106, for a particular recognized speech string, the recognition server will output a corresponding response signal to the voice response output 402 over the response select 406. The response select 406 will identify which one of a plurality of stored digital voice responses are to be announced. The corresponding voice string is accessed from the selected context voice response file 404 and played back by the voice response output.
>
> [Stanford, col. 11, line 66 through col. 12, line 15]

The operation described in the cited portion of Stanford does not teach or suggest the limitations of claim 14 in general nor the specific limitation of "(G) converting, in said fourth server object, said syntactic data message to a semantic data message that includes a plurality of possible valid interpretations of said voice data message."

The secondary reference, Christensen, is cited for teaching a system for remote objects to communicate over a computer network. Cristensen is therefore not understood to cure the deficiencies noted above for Stanford.

Thus, the combination of Stanford together with Christensen does not teach or suggest at least two limitations of amended independent method claim 14 (from which claims 15-18, 22, and 23 depend), i.e., "(C) converting said voice data message to a phonetic data message in said second server object," and "(G) converting, in said fourth server object, said syntactic data message to a semantic data message that includes a plurality of possible valid interpretations of said voice data message."

Accordingly, the teachings of the Stanford and Christensen references, either alone or in combination, fail to teach or suggest all of the limitations as arranged in claim 14. The references are therefore an improper basis for a rejection of claim 14 under 35 U.S.C. § 103(a) and claim 14 is patentable over these references. Because claims 15-18, 22, and 23 depend from claim 14 they are patentable for at least the same reason(s) as claim 14. The rejection should be withdrawn,

accordingly.

Claims 1-13, 19-21, and 24-38

Concerning item 6 of the Office Action, claims 1-13, 19-21, and 24-38 have been rejected under 35 U.S.C. § 103(a) as being unpatentable over Stanford taken in view of Christensen and further in view of U.S. Patent No. 5,675,723 to Ekrot et al. ("Ekrot"). Applicant notes that claims 6 and 29 were canceled by the previously-entered amendment. This rejection is respectfully traversed and reconsideration is requested for the following reasons.

As noted above, one requirement for a rejection under 35 U.S.C. § 103 is that the cited reference(s) teach or suggest all of the limitations of the claims at issue. In this situation, the combination of Stanford, Christensen, and Ekrot fails to teach or suggest all of the limitations of amended independent claims 1, 14, and 24, which are the base claims of the remaining claims under rejection.

As was described previously for the rejection of claim 14, the combination of Stanford together with Christensen does not teach or suggest at least two limitations of amended independent method claim 14 (from which claims 19-21 depend), i.e., "(C) converting said voice data message to a phonetic data message in said second server object," and "(G) converting, in said fourth server object, said syntactic data message to a semantic data message that includes a plurality of possible valid interpretations of said voice data message." Amended independent claims 1 and 24 contain similar limitations corresponding to those of method claim 14, i.e., "converting said voice data message to a phonetic data message," and "converting, in said fourth server object, said syntactic data message to a semantic data message that includes a plurality of possible valid interpretations of said voice data message."

Ekrot teaches a primary server/standby server network configuration that includes a primary server executing network operating system software and a standby server monitoring for the proper operation of the primary server, where both the primary server and the standby server are connected

to a storage medium. *See* Ekrot, Abstract. Ekrot is cited by the Examiner as teaching such a control monitor. Applicant, therefore, respectfully submits that Ekrot fails to cure the deficiencies noted above for Stanford and Christensen.

Consequently, the teachings of Ekrot, Stanford, and Christensen, whether considered alone or in combination, do not teach or suggest all of the limitations of independent claims 1, 14, and 24, which are therefore patentable over the combination of references. Because claims 2-5 and 7-13 depend from claim 1, either directly or indirectly, they are patentable for at least the same reason(s) as claim 1. Likewise, because claims 19-21 depend indirectly from claim 14, they are patentable for at least the same reason(s) as claim 14. Similarly, because claims 25-29 and 30-37 depend directly or indirectly from claim 24, they are patentable for at least the same reason(s) as claim 24. Thus, the combination of Ekrot, Stanford, and Christensen does not form a proper basis for a rejection of claims 1-13, 19-21, and 24-38 under 35 U.S.C. § 103(a). The rejection should be withdrawn, accordingly.

Conclusion

In view of the amendments and remarks submitted herein, Applicant respectfully submits that all of the claims now pending in the subject application are in condition for allowance, and respectfully requests a Notice of Allowance for the application.

Authorization is hereby given to charge a three-month extension of time under 37 C.F.R. § 1.136 or any other required fees and/or to credit any overpayments to deposit account No. 50-1133.

If the Examiner believes there are any outstanding issues to be resolved with respect to the above-identified application, the Examiner is invited to telephone the undersigned at his earliest convenience so that such issues may be resolved telephonically.

Respectfully submitted,

Date: 3.13.06

Toby H. Kusmer, Reg. No. 26,418
G. Matthew McCloskey, Reg. No. 47,025
McDermott Will & Emery LLP
28 State Street
Boston, MA 02109
V: (617) 535-4065
F: (617) 535-3800